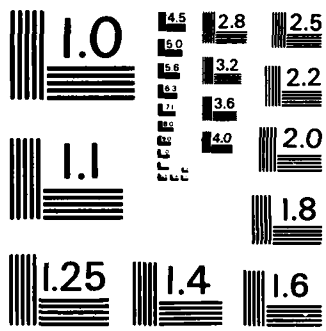
AD-A159 059 INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARI..(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS B K SINHA ET AL. JUL 85 TR-85-27 1/1

UNCLASSIFIED AFOSR-TR-85-0698 F49620-85-C-0008 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 85-0698

Center for Multivariate Analysis

University of Pittsburgh

2

INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX AND THE GENERALIZED VARIANCE UNDER ENTROPY LOSS

Bimal K. Sinha[1]
University of Pittsburgh

and

Malay Ghosh[2]
University of Florida

July 1985

Technical Report No. 85-27

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

[1]Research partially supported by the Air Force Office of Scientific Research (AFOSR), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

[2]Research partially supported by the NSF Grant Number DMS-8218191 and partially by the Faculty Development award from the DSR, University of Florida.

INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX AND THE GENERALIZED VARIANCE UNDER ENTROPY LOSS

by

Bimal K. Sinha[1]
University of Pittsburgh

and

Malay Ghosh[2]
University of Florida

A-1

Abstract

Based on a data matrix $X = (\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_k)$: $p \times k$ with independent columns $\underset{\sim}{X}_i \sim N_p(\underset{\sim}{\xi}_i, \Sigma)$, and an independent Wishart matrix S: $p \times p \sim W_p(n, \Sigma)$, testimators dominating the best equivariant estimators of Σ and $|\Sigma|$ are obtained under two types of entropy loss. For simultaneous estimation of the mean vector and the variance covariance matrix of a multinormal population, a suitable entropy loss is developed and testimators dominating the pair consisting of the sample mean vector and the best multiple of the sample Wishart matrix are derived. A technique of SINHA (Jour. Mult. Analysis, 1976) is heavily exploited.

AMS 1980 Subject Classification. Primary 62F10; Secondary 62H99

Key Words: Best equivariant estimator, entropy loss, generalized variance, MANOVA test, Roy's maximum root test, testimator, Wishart distribution.

[1] Research partially supported by the Air Force Office of Scientific Research (AFOSR), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

[2] Research partially supported by the NSF Grant Number DMS-8218191 and partially by the Faculty Development award from the DSR, University of Florida.

INADMISSIBILITY OF THE BEST EQUIVARIANT ESTIMATORS OF THE VARIANCE-COVARIANCE MATRIX AND THE GENERALIZED VARIANCE UNDER ENTROPY LOSS

by

Bimal K. Sinha[1] and Malay Ghosh[2]

University of Pittsburgh and University of Florida

1. INTRODUCTION. Suppose $Y_1,\ldots,Y_n$ are iid $N(\xi,\sigma^2)$. If ξ is known, then the best scale invariant estimator of σ^2 is given by

$$\phi_0(Y_1,\ldots,Y_n) = (n+2)^{-1}\sum_{i=1}^n (Y_i - \xi)^2. \tag{1.1}$$

It is proved in Girshick and Savage (1951), and Hodges and Lehmann (1951) that ϕ_0 is an admissible estimator of σ^2 under squared error loss. However, if ξ is unknown, then Stein (1964) has shown that the natural estimator

$$\phi_1(Y_1,\ldots,Y_n) = (n+1)^{-1}\sum_{i=1}^n (Y_i - \bar{Y})^2, \tag{1.2}$$

of $\sigma^2 (\bar{Y} = n^{-1}\sum_{i=1}^n Y_i)$ is inadmissible under squared error loss, and is dominated by estimators of the form

$$\phi(Y_1,\ldots,Y_n) = \min[(n+1)^{-1}\sum_{i=1}^n (Y_i - \bar{Y})^2, (n+2)^{-1}\sum_{i=1}^n (Y_i - \xi_0)^2] \tag{1.3}$$

for every fixed constant ξ_0. The estimator ϕ of σ^2 can be viewed as a preliminary test estimator (testimator) which uses the estimator $(n+1)^{-1}\sum_{i=1}^n (Y_i-\bar{Y})^2$ for σ^2 if the F-statistic $n(\bar{Y} - \xi_0)^2/\{(n-1)^{-1}\sum_{i=1}^n (Y_i-\bar{Y})^2\}$ for testing H_0: $\xi = \xi_0$ against the alternatives H_1: $\xi \neq \xi_0$ exceeds $(n-1)/(n+1)$ (thereby rejecting H_0 at a certain

---

[1]Research partially supported by the Air Force Office of Scientific Research (AFOSR), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

[2]Research partially supported by the NSF Grant Number DMS-8218191 and partially by the Faculty Development award from the DSR, University of Florida.

significance level), and uses the estimator $(n+2)^{-1}\sum_{i=1}^{n}(Y_i-\xi_0)^2$ otherwise. Stein (1964) considered the more general regression analog of this problem in the canonical set up. Brewster and Zidek (1974) have shown that the results extend to a more general loss including the entropy loss (first introduced in James and Stein (1961)) given by

$$L(a,\sigma^2) = a/\sigma^2 - \log(a/\sigma^2) - 1, \tag{1.4}$$

to which attention will be restricted in this paper.

There are two possible multivariate extensions of the above results. One can consider estimation of the variance-covariance matrix $\underset{\sim}{\Sigma}$ or the generalized variance $|\underset{\sim}{\Sigma}|$ in a multinormal set up. To fix ideas, let $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_m$ be iid $N(\underset{\sim}{\xi},\underset{\sim}{\Sigma})$, where each $\underset{\sim}{Y}_i$ is $p\times 1$. When both $\underset{\sim}{\xi}$ and $\underset{\sim}{\Sigma}$ are unknown, the minimal sufficient statistic for these parameters is $(\underset{\sim}{X},\underset{\sim}{S})$, where $\underset{\sim}{X} = m^{1/2}\bar{\underset{\sim}{Y}}_m$ and $\underset{\sim}{S} = \sum_{i=1}^{m}(\underset{\sim}{Y}_i-\bar{\underset{\sim}{Y}}_m)(\underset{\sim}{Y}_i-\bar{\underset{\sim}{Y}}_m)^T$ $(\bar{\underset{\sim}{Y}}_m = m^{-1}\sum_{i=1}^{m}\underset{\sim}{Y}_i)$. Haff (1979b, 1980, 1982) and Dey and Srinivasan (1985) have considered estimation of $\underset{\sim}{\Sigma}$ under several losses including the entropy loss

$$L_1(\hat{\underset{\sim}{\Sigma}},\underset{\sim}{\Sigma}) = \mathrm{tr}(\hat{\underset{\sim}{\Sigma}}\underset{\sim}{\Sigma}^{-1}) - \log|\hat{\underset{\sim}{\Sigma}}\underset{\sim}{\Sigma}^{-1}| - p. \tag{1.5}$$

They propose estimators $\hat{\underset{\sim}{\Sigma}}$ of $\underset{\sim}{\Sigma}$ which are functions of the Wishart matrix $\underset{\sim}{S}$ alone, but do not consider any Stein-type estimators (i.e. testimators).

In this note, we consider estimation of $\underset{\sim}{\Sigma}$ and $\underset{\sim}{\Sigma}^{-1}$ each under the entropy losses L_1 and

$$L_2(\hat{\underset{\sim}{\Sigma}},\underset{\sim}{\Sigma}) = \mathrm{tr}(\underset{\sim}{\Sigma}\hat{\underset{\sim}{\Sigma}}^{-1}) - \log|\underset{\sim}{\Sigma}\hat{\underset{\sim}{\Sigma}}^{-1}| - p. \tag{1.6}$$

To our knowledge, the loss (1.6) has never been considered before either for estimating $\underset{\sim}{\Sigma}$ or $\underset{\sim}{\Sigma}^{-1}$. Haff (1977, 1979a, 1979b) considers estimation of $\underset{\sim}{\Sigma}^{-1}$ under various quadratic losses. For us, the loss (1.6) seems to be as natural as (1.5), and can be motivated as follows. Suppose $\underset{\sim}{S}$ is a random variable having a p-dimensional Wishart distribution with degrees of freedom n and parameter $\underset{\sim}{\Sigma}$(to be denoted by $W_p(n,\Sigma)$). Write $f_{\Sigma}(S)$ as the pdf of S. Then, a meaningful loss in estimating Σ by A (or Σ^{-1} by A^{-1}) is the entropy distance between $W_p(n,\Sigma)$ and $W_p(n,A)$, and is given by

$$E_{\Sigma}[\log \frac{f_{\Sigma}(S)}{f_A(S)}] = (n/2)L_2(A,\Sigma). \tag{1.7}$$

Use of an estimator $\hat{\Sigma}$ in place of A gives rise to (1.6).

In Section 2 of this note, we consider estimation of Σ and Σ^{-1} each under the losses (1.5) and (1.6), and develop Stein-type testimators dominating the best multiples of the Wishart matrix and its inverse. A technique of Sinha (1976) is heavily exploited. Incidentally, it may be remarked that for the loss (1.6) no Haff-type improved estimator over the best equivariant estimator is readily available. We also consider simultaneous estimation of the mean vector and the variance-covariance matrix, and develop certain estimators dominating the pair consisting of the sample mean vector and the best multiple of the Wishart matrix under a suitable entropy loss to be developed in Section 2.

In Section 3, we consider estimation of $|\Sigma|$. This problem has received attention in Shorrock and Zidek (1976), and Sinha (1976). In these papers Stein-type testimators are developed, and are shown to dominate the best multiple of $|S|$ under squared error loss. Similar testimators are developed in Section 3, and are shown to dominate the best multiple of $|S|$ under the two entropy losses

$$L_1(|\hat{\Sigma}|,|\Sigma|) = |\hat{\Sigma}|/|\Sigma| - \log(|\hat{\Sigma}|/|\Sigma|) - 1 \tag{1.8}$$

and

$$L_2(|\hat{\Sigma}|,|\Sigma|) = |\Sigma|/|\hat{\Sigma}| - \log(|\Sigma|/|\hat{\Sigma}|) - 1. \tag{1.9}$$

Throughout this paper, for two matrices A and B of the same order, $A \geq B$ implies that $A - B$ is nonnegative definite. In the remainder of this section, we state with-out proof three matrix lemmas which are used repeatedly in Section 2. The proofs of these lemmas are quite straightforward.

LEMMA 1. *Let $\mathcal{F}_p$ denote the class of all nonsingular matrices. Then for any* $A \in \mathcal{F}_p$ and $B \in \mathcal{F}_p$,

$$\mathrm{tr}(AB) - \log|B| \geq \log|A| + p, \tag{1.10}$$

equality holding iff $B = A^{-1}$.

LEMMA 2. *Suppose* $A \geq 0$ *and* $B \geq C$, *where* A, B, C *and the null matrix* 0 *are square matrices of the same order. Then*,

$$\mathrm{tr}\, AB \geq \mathrm{tr}\, AC. \tag{1.11}$$

LEMMA 3. *For any positive definite matrix* A,

$$\mathrm{tr}\, A - \log|A| - p \geq 0,$$

with equality iff $A = I_p$.

2. ESTIMATION OF Σ AND Σ^{-1}. Consider a multivariate normal linear model in its canonical form. Suppose $X = (X_1, \ldots, X_k)$ is a $p\times k$ matrix with independent columns $X_i \sim N_p(\xi_i, \Sigma)$, and let S be a p-dimensional Wishart matrix with degrees of freedom n and parameter Σ distributed independently of X. We assume $n > p+1$ and ξ_i's unknown.

Consider first estimation of Σ under the loss (1.5). As pointed out by Shorrock and Zidek (1976), the above problem remains invariant under the full affine group G acting on the space of $p\times k$ matrices (writing $\xi = (\xi_1, \ldots, \xi_k)$)

$$X \to AX + B,\ \xi \to A\xi + B,\ S \to ASA^T,\ \Sigma \to A\Sigma A^T, \tag{2.1}$$

where $\underset{\sim}{A}$ is any nonsingular p×p matrix, and B is any p×k matrix. Then, any affine equivariant estimator of $\underset{\sim}{\Sigma}$ must be of the form

$$\phi_0(\underset{\sim}{S}) = cS, \tag{2.2}$$

where c is a constant. Noting that $E(\underset{\sim}{S}) = n$, it follows that under the loss (1.5), the optimal choice of c minimizing the risk of $c\underset{\sim}{S}$ under the loss (1.5) is $c = n^{-1}$.

Following Sinha (1976), write $\underset{\sim}{S} = \underset{\sim}{W}\underset{\sim}{W}^T$ and $\underset{\sim}{U} = \underset{\sim}{W}^{-1}\underset{\sim}{X}$, where $\underset{\sim}{W}$ is a p×p nonsingular matrix. In order to improve on the best affine equivariant estimator $n^{-1}\underset{\sim}{S}$, consider the class C of estimators of $\underset{\sim}{\Sigma}$ having the form $\hat{\underset{\sim}{\Sigma}} = \underset{\sim}{\phi}(\underset{\sim}{W},\underset{\sim}{U}) = \underset{\sim}{W}\underset{\sim}{\psi}\underset{\sim}{W}^T$, where $\underset{\sim}{\psi} \equiv \underset{\sim}{\psi}(\underset{\sim}{U}\underset{\sim}{U}^T)$ is a p×p nonsingular matrix. This class C contains estimators equivariant under a nonnormal subgroup H of G obtained from G by putting $B = \underset{\sim}{0}$. The special choice $\underset{\sim}{\psi} \equiv \underset{\sim}{\psi}_0 = n^{-1}\underset{\sim}{I}_p$ leads to the corresponding estimator $\underset{\sim}{\phi}_0 = \underset{\sim}{W}\underset{\sim}{\psi}_0\underset{\sim}{W}^T = n^{-1}\underset{\sim}{S}$ of $\underset{\sim}{\Sigma}$. In order to compute the risk $R_{\underset{\sim}{\phi}}$ of $\underset{\sim}{\phi}$ under the loss (1.5), first let $\underset{\sim}{X}_* = \underset{\sim}{A}\underset{\sim}{X}$, $\underset{\sim}{W}_* = \underset{\sim}{A}\underset{\sim}{W}$, where $\underset{\sim}{A}$ is a p×p nonsingular matrix such that $\underset{\sim}{A}\underset{\sim}{\Sigma}\underset{\sim}{A}^T = \underset{\sim}{I}_p$, $\underset{\sim}{I}_p$ denoting the identity matrix of order p. Note that $\underset{\sim}{U} = \underset{\sim}{W}^{-1}\underset{\sim}{X} = \underset{\sim}{W}_*^{-1}\underset{\sim}{X}_*$. Then, writing $E_{\xi,\Sigma}$ as the expectation under $N_p(\underset{\sim}{\xi}_i,\underset{\sim}{\Sigma})$ for $\underset{\sim}{X}_i$ $i = 1,\ldots,p$ and $W_p(n,\Sigma)$ for S, and $\underset{\sim}{\xi}_* = \underset{\sim}{A}\underset{\sim}{\xi}$, one gets

$$\begin{aligned} R_{\underset{\sim}{\phi}} &= E_{\underset{\sim}{\xi},\underset{\sim}{\Sigma}}[\operatorname{tr}(\underset{\sim}{W}\underset{\sim}{\psi}\underset{\sim}{W}^T\underset{\sim}{\Sigma}^{-1}) - \log|\underset{\sim}{W}\underset{\sim}{\psi}\underset{\sim}{W}^T\underset{\sim}{\Sigma}^{-1}| - p] \\ &= E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[\operatorname{tr}(\underset{\sim}{W}_*^T\underset{\sim}{W}_*\underset{\sim}{\psi}) - \log|W_*^TW_*| - \log|\underset{\sim}{\psi}| - p]. \end{aligned} \tag{2.3}$$

Note that for comparing the risk performance of members within the class $C \ni \underset{\sim}{\phi} = \underset{\sim}{W}\underset{\sim}{\psi}\underset{\sim}{W}^T$ (under $\underset{\sim}{\xi},\underset{\sim}{\Sigma}$) $= \underset{\sim}{W}_*\underset{\sim}{\psi}\underset{\sim}{W}_*^T$ (under $\underset{\sim}{\xi}_*,I_p$), it suffices to consider

$$\begin{aligned} \tilde{R}_{\underset{\sim}{\phi}} &= E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[\operatorname{tr}(\underset{\sim}{W}_*^T\underset{\sim}{W}_*\underset{\sim}{\psi}) - \log|\underset{\sim}{\psi}|] \\ &= \tilde{E}[\operatorname{tr}\{\underset{\sim}{\psi}_*^{-1}(\underset{\sim}{\xi}_*,\underset{\sim}{U})\underset{\sim}{\psi}\} - \log|\underset{\sim}{\psi}|], \end{aligned} \tag{2.4}$$

where $\psi_*^{-1}(\xi_*,u) = E_{\xi_*,I_p}(W_*^T W_* | U = u)$, and $\tilde{E}$ denotes expectation over the marginal distribution of U.

To minimize R_ϕ with respect to ϕ, it suffices to minimize

$$\mathrm{tr}\{\psi_*^{-1}(\xi_*,u)\psi\} - \log|\psi| \tag{2.5}$$

with respect to ψ for every u. Using Lemma 1 with $A = \psi_*^{-1}(\xi_*,u)$ and $B = \psi$, it follows that the expression in (2.5) is minimized when $\psi = \psi_*(\xi_*,u)$. However, this expression involves not only u but ξ_* also. We find next an upper bound for $\psi_*(\xi_*,u)$ free from ξ_*.

<u>LEMMA 4</u>. $\psi_*(\xi_*,u) \leq \tilde{\psi}(u) = (n+k)^{-1}(I_p + uu^T)$.

The proof of this lemma is deferred to the Appendix. Based on $\tilde{\psi}(u)$, a testimator is now constructed as follows.

Let $\tilde{\phi} = W\tilde{\psi}W^T = (n+k)^{-1}(S+XX^T)$. For estimating Σ, define the testimator

$$\begin{aligned} \tilde{\tilde{\phi}} &= \tilde{\phi} \quad \text{if } \tilde{\phi} \leq \phi_0 \\ &= \phi_0 \quad \text{otherwise.} \end{aligned} \tag{2.6}$$

The corresponding ψ say $\tilde{\tilde{\psi}}$ is given by

$$\begin{aligned} \tilde{\tilde{\psi}} &= \tilde{\psi} \text{ if } \tilde{\psi} \leq \psi_0 \\ &= \psi_0 \text{ otherwise.} \end{aligned} \tag{2.7}$$

<u>Remark 1</u>. The estimator $\tilde{\tilde{\phi}}$ defined in (2.6) is a multivariate generalization of Stein's (1964) univariate testimator. The condition $\tilde{\phi} \leq \phi_0$ can be alternately expressed as $XX^T \leq (k/n)S \iff \sup_{\ell \neq 0}(\ell^T XX^T \ell)/(\ell^T S \ell) \leq k/n \iff$ largest eigenvalue of $X^T S^{-1} X \leq k/n$. Thus, the estimator proposed in (2.6) is based on Roy's maximum root

test. The test reduces for $k=1$ to Hotelling's T^2 test.

Remark 2. The condition $\tilde{\psi} \le \psi_0$ can be alternately expressed as $UU^T \le k/n \iff \|U\| \le k/n$ where $\|\cdot\|$ denotes the Euclidean norm.

The following theorem shows the dominance of the testimator $\tilde{\tilde{\phi}}$ over ϕ_0.

THEOREM 1. Under the loss (1.5),

$$R_{\tilde{\tilde{\phi}}} < R_{\phi_0} \quad \text{for all } \xi \text{ and } \Sigma.$$

Proof: Using (2.3), (2.4) and (2.7),

$$R_{\tilde{\tilde{\phi}}} - R_{\phi_0} = \tilde{R}_{\tilde{\tilde{\phi}}} - \tilde{R}_{\phi_0}$$

$$= \tilde{E}[\operatorname{tr}(\psi_*^{-1}(\tilde{\psi} - \psi_0)) - \log \frac{|\tilde{\psi}|}{|\psi_0|}] I_{[\tilde{\psi} \le \psi_0]} \ . \tag{2.8}$$

From Lemma 4, $\psi_*^{-1} \ge \tilde{\psi}^{-1}$. Now, using Lemma 2 with $A = \psi_0 - \tilde{\psi}$, $B = \psi_*^{-1}$ and $C = \tilde{\psi}^{-1}$, one gets from (2.8),

$$R_{\tilde{\tilde{\phi}}} - R_{\phi_0} \le -\tilde{E}[\operatorname{tr}\tilde{\psi}^{-1}(\psi_0 - \tilde{\psi}) + \log \frac{|\tilde{\psi}|}{|\psi_0|}] I_{[\tilde{\psi} \le \psi_0]}$$

$$= -\tilde{E}[\operatorname{tr}(\psi_0 \tilde{\psi}^{-1}) - \log|\psi_0 \tilde{\psi}^{-1}| - p] I_{[\tilde{\psi} \le \psi_0]}$$

$$< 0, \tag{2.9}$$

where in the last step of (2.9), one uses Lemma 3. The proof of Theorem 1 is complete from (2.8) and (2.9).

Remark 3. Quite generally, given any estimator $\phi = W\psi W^T$, defining $\tilde{\tilde{\phi}} = \tilde{\phi}$ if $\tilde{\phi} \le \phi$, $= \phi$ otherwise, where $\tilde{\phi} = W\tilde{\psi}W^T$ and $\tilde{\psi}$ is defined in Lemma 4, one gets $R_{\tilde{\tilde{\phi}}} < R_{\phi}$ for all ξ and Σ (vide Sinha (1976)). This enables one to develop sequential testimators as in Sinha (1976).

Next we consider the loss given in (1.6). In this case, the best multiple of $\underset{\sim}{S}$ (minimizing the risk) is given by $(n-p-1)^{-1}$. Let $\phi_{00}(\underset{\sim}{S}) = (n-p-1)^{-1}\underset{\sim}{S}$, $= \underset{\sim}{W}\underset{\sim}{\psi}_{00}\underset{\sim}{W}^T$ so that $\underset{\sim}{\psi}_{00} = (n-p-1)^{-1}\underset{\sim}{I}_p$. Once again, we consider a competing class C_0 of estimators of the form $\underset{\sim}{\phi}(\underset{\sim}{S}) = \underset{\sim}{W}\underset{\sim}{\psi}(\underset{\sim}{U}\underset{\sim}{U}^T)\underset{\sim}{W}^T$. Proceeding as in (2.3), under the loss (1.6), the risk of $\underset{\sim}{\phi}$ is given by

$$R_{\underset{\sim}{\phi}} = E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[\mathrm{tr}\{(\underset{\sim}{W}_*^T\underset{\sim}{W}_*)^{-1}\underset{\sim}{\psi}^{-1}\} - \log|(\underset{\sim}{W}_*^T\underset{\sim}{W}_*)^{-1}| - \log|\underset{\sim}{\psi}^{-1}| - p].$$

Hence, for comparing estimators of the given type $\underset{\sim}{\phi}$, it suffices to consider

$$\tilde{R}_{\underset{\sim}{\phi}} = E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[\mathrm{tr}\{(\underset{\sim}{W}^T\underset{\sim}{W}_*)^{-1}\underset{\sim}{\psi}^{-1}\} - \log|\underset{\sim}{\psi}^{-1}|]. \tag{2.10}$$

Using Lemma 1 once again, it follows that the optimal choice of $\underset{\sim}{\psi}$ is $E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[(\underset{\sim}{W}_*^T\underset{\sim}{W}_*)^{-1}|\underset{\sim}{u}]$ $= \underset{\sim}{\psi}_1(\underset{\sim}{\xi}_*,\underset{\sim}{u})$ (say). Similar to Lemma 4, we now prove the following lemma.

<u>LEMMA 5</u>. $\underset{\sim}{\psi}_1(\underset{\sim}{\xi}_*,\underset{\sim}{u}) \le (n-p-1+k)^{-1}(\underset{\sim}{I}_p+\underset{\sim}{u}\underset{\sim}{u}^T) = \tilde{\underset{\sim}{\psi}}_0(\underset{\sim}{u})$ (say).

The proof of Lemma 5 is also deferred to the appendix. Let $\tilde{\underset{\sim}{\phi}}_0(\underset{\sim}{S}) = \underset{\sim}{W}\tilde{\underset{\sim}{\psi}}_0(\underset{\sim}{u})\underset{\sim}{W}^T$. Similar to the previous situation, we define the testimator

$$\begin{aligned}\tilde{\tilde{\underset{\sim}{\phi}}}_0(\underset{\sim}{S}) &= \tilde{\underset{\sim}{\phi}}_0(\underset{\sim}{S}) \quad \text{if } \tilde{\underset{\sim}{\phi}}_0(\underset{\sim}{S}) \le \underset{\sim}{\phi}_{00}(\underset{\sim}{S}) \\ &= \underset{\sim}{\phi}_{00}(\underset{\sim}{S}) \quad \text{otherwise.}\end{aligned} \tag{2.11}$$

Accordingly, the $\underset{\sim}{\psi}$ value corresponding to $\tilde{\tilde{\underset{\sim}{\phi}}}_0$ will be given by

$$\begin{aligned}\tilde{\tilde{\underset{\sim}{\psi}}}_0(\underset{\sim}{u}) &= \tilde{\underset{\sim}{\psi}}_0(\underset{\sim}{u}) \quad \text{if } \tilde{\underset{\sim}{\psi}}_0(\underset{\sim}{u}) \le \psi_{00}(\underset{\sim}{u}) \\ &= \underset{\sim}{\psi}_{00}(\underset{\sim}{u}) \quad \text{otherwise.}\end{aligned} \tag{2.12}$$

<u>Remark 4</u>. Note that $\tilde{\underset{\sim}{\phi}}_0(\underset{\sim}{S}) = (n-p-1+k)^{-1}(\underset{\sim}{S}+\underset{\sim}{X}\underset{\sim}{X}^T)$. Hence, the condition $\tilde{\underset{\sim}{\phi}}_0(\underset{\sim}{S}) \le \phi_{00}(\underset{\sim}{S})$ can be equivalently expressed as largest eigenvalue of $\underset{\sim}{X}'\underset{\sim}{S}^{-1}\underset{\sim}{X} \le k/(n-p-1)$. Thus, in this case also, the preliminary test is based on Roy's maximum root test.

We now prove the following theorem.

THEOREM 2. Under the loss (1.6), $R_{\tilde{\phi}_0} < R_{\phi_{00}}$.

Proof: Write

$$R_{\tilde{\phi}_0} - R_{\phi_{00}} = \tilde{R}_{\tilde{\phi}_0} - \tilde{R}_{\phi_{00}}$$

$$= \tilde{E}[\mathrm{tr}\{\psi_1(\xi_*,U)(\tilde{\psi}_0^{-1} - \psi_{00}^{-1})\} - \log \frac{|\tilde{\psi}_0^{-1}|}{|\psi_{00}^{-1}|}] I_{[\tilde{\psi}_0 \leq \psi_{00}]} \tag{2.13}$$

In view of Lemma 5, putting $A = \tilde{\psi}_0^{-1} - \psi_{00}^{-1}$, $B = \tilde{\psi}_0$ and $C = \psi_1(\xi_*,U)$ in Lemma 2 one gets

$$\begin{aligned} &\text{rhs of (2.13)} \\ &\leq \tilde{E}[\mathrm{tr}(\tilde{\psi}_0(\tilde{\psi}_0^{-1} - \psi_{00}^{-1}) + \log|\tilde{\psi}_0\psi_{00}^{-1}|] I_{[\tilde{\psi}_0 \leq \psi_{00}]} \\ &= -\tilde{E}[\mathrm{tr}(\tilde{\psi}_0\psi_{00}^{-1}) - \log|\tilde{\psi}_0\psi_{00}^{-1}| - p] I_{[\tilde{\psi}_0 \leq \psi_{00}]} \\ &< 0, \end{aligned} \tag{2.14}$$

where in the last step of (2.14), one uses Lemma 3. The proof of the theorem is complete from (2.13) and (2.14).

Remark 5. Here again, as explained in Remark 3, one can develop sequential testimators each dominating the best equivariant estimator $\phi_{00}(S)$.

We consider now the simultaneous estimation of the mean vector and the variance-covariance matrix under entropy loss. Writing $f_{\mu,\Sigma}(Y_1,\ldots,Y_n)$ as the joint pdf of n iid $N_p(\mu,\Sigma)$ variables, it follows that taking the loss $L((\mu,\Sigma), (\ell,A))$ as the entropy distance between the $N_p(\mu,\Sigma)$ and $N_p(\ell,A)$ distributions, and assuming $n > p+1$ for purposes of estimation, one gets

$$L[(\underset{\sim}{\mu},\underset{\sim}{\Sigma}),(\underset{\sim}{\ell},\underset{\sim}{A})] = E_{\underset{\sim}{\mu},\underset{\sim}{\Sigma}}[\log \frac{f_{\underset{\sim}{\mu},\underset{\sim}{\Sigma}}(\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n)}{f_{\underset{\sim}{\ell},\underset{\sim}{A}}(\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n)}]$$

$$= E_{\underset{\sim}{\mu},\underset{\sim}{\Sigma}}[\log(|\underset{\sim}{\Sigma}|/|\underset{\sim}{A}|)^{-\frac{n}{2}} - \frac{1}{2}\{tr(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{S}_0) + n(\bar{\underset{\sim}{Y}}-\underset{\sim}{\mu})^T\underset{\sim}{\Sigma}^{-1}(\bar{\underset{\sim}{Y}}-\underset{\sim}{\mu})\}$$

$$+ \frac{1}{2}\{tr(\underset{\sim}{A}^{-1}\underset{\sim}{S}_0) + n(\bar{\underset{\sim}{Y}}-\underset{\sim}{\ell})^T\underset{\sim}{\Sigma}^{-1}(\bar{\underset{\sim}{Y}}-\underset{\sim}{\ell})\}]$$

(where $\bar{\underset{\sim}{Y}} = n^{-1}\sum_{i=1}^n \underset{\sim}{Y}_i$, $\underset{\sim}{S}_0 = \sum_{i=1}^n (\underset{\sim}{Y}_i-\bar{\underset{\sim}{Y}})(\underset{\sim}{Y}_i-\bar{\underset{\sim}{Y}})^T$)

$$= \log|\underset{\sim}{\Sigma}\underset{\sim}{A}^{-1}|^{-\frac{n}{2}} - \frac{np}{2} - \frac{p}{2}$$

$$+ \frac{n}{2}\, tr(\underset{\sim}{\Sigma}\underset{\sim}{A}^{-1}) + \frac{1}{2}[p + n(\underset{\sim}{\ell}-\underset{\sim}{\mu})^T\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\ell}-\underset{\sim}{\mu})]$$

$$= \frac{n}{2}[(\underset{\sim}{\ell}-\underset{\sim}{\mu})^T\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\ell}-\underset{\sim}{\mu}) + tr(\underset{\sim}{\Sigma}\underset{\sim}{A}^{-1}) - \log|\underset{\sim}{\Sigma}A^{-1}| - p]. \tag{2.15}$$

Thus if one uses the best location and scale invariant estimators $\bar{\underset{\sim}{Y}}$ and $(n-p-1)^{-1}\underset{\sim}{S}_0$ for $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$ respectively, in view of the loss (2.15), it suffices to improve on $\bar{\underset{\sim}{Y}}$ and $(n-p-1)^{-1}\underset{\sim}{S}_0$ separately. From James and Stein (1961), it follows that under the loss

$$L_0(\underset{\sim}{\ell},\underset{\sim}{\mu}) = (\underset{\sim}{\ell}-\underset{\sim}{\mu})^T\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\ell}-\underset{\sim}{\mu}),$$

$\bar{\underset{\sim}{Y}}$ is improved by the estimator $\underset{\sim}{\delta}_0(\bar{\underset{\sim}{Y}},\underset{\sim}{S}_0) = (1 - \frac{p-2}{n\bar{\underset{\sim}{Y}}^T\underset{\sim}{S}_0^{-1}\bar{\underset{\sim}{Y}}})\bar{Y}$, provided $p > 2$.
Again the loss

$$L_2(\underset{\sim}{\Sigma},\underset{\sim}{A}) = tr(\underset{\sim}{\Sigma}\underset{\sim}{A}^{-1}) - \log|\underset{\sim}{\Sigma}\underset{\sim}{A}^{-1}| - p$$

is the same as the loss (1.6), and taking $k = 1$ in Theorem 2, it follows that $(n-p-1)^{-1}\underset{\sim}{S}_0$ is dominated by the testimator

$$\underset{\sim}{\delta}_1(\bar{\underset{\sim}{Y}},\underset{\sim}{S}_0) = (n-p)^{-1}(\underset{\sim}{S}_0+n\bar{\underset{\sim}{Y}}\bar{\underset{\sim}{Y}}^T) \quad \text{if } (n-p)^{-1}(\underset{\sim}{S}_0+n\bar{\underset{\sim}{Y}}\bar{\underset{\sim}{Y}}^T) \le (n-p-1)^{-1}\underset{\sim}{S}_0$$

$$= (n-p-1)^{-1}\underset{\sim}{S}_0 \quad \text{otherwise.}$$

Thus, under the loss (2.15), $(\bar{Y}, S_0)$ is dominated by (δ_0, δ_1) for $p > 2$. For $p \leq 2$, $(\bar{Y}, S_0)$ is of course dominated by $(\bar{Y}, \delta_1)$. Finally, for $p \geq 2$, sequential testimators each dominating $(\bar{Y}, S_0)$ for the loss (2.15) can be easily obtained (vide Sinha (1976)).

Next we consider estimation of the precision matrix Σ^{-1} under the losses (1.5) and (1.6) (calling $L_1(\hat{\Sigma}, \Sigma)$ and $L_2(\hat{\Sigma}, \Sigma)$ as $L_1(\hat{\Sigma}^{-1}, \Sigma^{-1})$ and $L_2(\hat{\Sigma}^{-1}, \Sigma^{-1})$ respectively). Consider the class of estimators of the form $\phi = (W\psi W^T)^{-1}$ for Σ^{-1} so that, under L_1, the choice $\psi = n^{-1} I_p$ leads to the best multiple (of S^{-1}) estimator nS^{-1} of Σ^{-1}. Recalling that $\tilde{\phi}(S) = (n+k)^{-1}(S+XX^T)$, it follows that defining $\tilde{\tilde{\phi}}$ as in (2.6), the best equivariant estimator $nS^{-1} = \phi_0^{-1}(S)$ is dominated by $\tilde{\tilde{\phi}}^{-1}(S)$. Similarly, under the loss L_2, the best equivariant estimator $(n-p-1)S^{-1} = \phi_{00}^{-1}(S)$ of Σ^{-1} is dominated by $\tilde{\tilde{\phi}}_0^{-1}(S)$ defined in (2.11). Again, in each case sequential testimators are easily obtained.

3. ESTIMATION OF $|\Sigma|$. Consider the same set up as of Section 2. Estimating $|\Sigma|$ by a, assume the loss to be given by

$$L_1(a, |\Sigma|) = \frac{a}{|\Sigma|} - \log \frac{a}{|\Sigma|} - 1. \tag{3.1}$$

Following Shorrock and Zidek (1976) and Sinha (1976), it follows that the best equivariant estimator of $|\Sigma|$ is $c_0|S|$ where c_0 is determined from minimizing $E_{\Sigma = I_p}$ $(c|S| - \log c)$ with respect to c. This gives

$$c_0 = (E_{\Sigma = I_p} |S|)^{-1} = (n-p)!/n!\ .$$

Following Stein's suggestion, and arguments as in Shorrock and Zidek (1976) or Sinha (1976), we look for better estimators in the class $\phi(X, S) = \psi(X^T S^{-1} X)|S|$

for some real valued function ψ. Under the loss (3.1), ϕ has the risk

$$R_\phi = E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[\psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X})|\underset{\sim}{S}| - \log\psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X}) - \log|\underset{\sim}{S}| - 1], \tag{3.2}$$

where $\underset{\sim}{\xi}_*$ is defined as in Section 2. Write $\underset{\sim}{V} = \underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X}$ so that given $\underset{\sim}{V} = \underset{\sim}{v}$, the best choice of ψ (minimizing (3.2)) is given by $\psi(\underset{\sim}{v}) = \psi_{\underset{\sim}{\xi}_*}(\underset{\sim}{v}) = \{E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}(|\underset{\sim}{S}|\,|\underset{\sim}{V}=\underset{\sim}{v})\}^{-1}$. Following the line of argument of Sinha (1976), one can easily show that

$$\psi_{\underset{\sim}{\xi}_*}(\underset{\sim}{v}) \le |\underset{\sim}{I}_k+\underset{\sim}{v}|(n-p+k)!/(n+k)! = \psi_0(\underset{\sim}{v}) \quad \text{(say)}. \tag{3.3}$$

Then, it is easy to show using (3.3) and strict convexity of the loss (3.1) that for every ψ defining $\tilde{\psi}(\underset{\sim}{v}) = \min(\psi(\underset{\sim}{v}),\psi_0(\underset{\sim}{v}))$, $\tilde{\psi}(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X})|\underset{\sim}{S}|$ dominates $\psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X})|\underset{\sim}{S}|$ under the loss (3.1). In particular, the estimator $Z = \min\{\frac{(n-p)!}{n!}|\underset{\sim}{S}|, \frac{(n-p+k)!}{(n+k)!}|\underset{\sim}{S}+\underset{\sim}{X}\underset{\sim}{X}^T|\}$ dominates $\{(n-p)!/n!\}|\underset{\sim}{S}|$ under the loss (3.1). Note that Z is indeed a testimator since the ratio $|\underset{\sim}{S}+\underset{\sim}{X}\underset{\sim}{X}^T|/|\underset{\sim}{S}| = |\underset{\sim}{I}_k+\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X}|$ is a MANOVA test statistic for H_0: $\underset{\sim}{\xi} = \underset{\sim}{0}$ against H_1: $\underset{\sim}{\xi} \ne \underset{\sim}{0}$.

For the other loss $L_2(a,|\underset{\sim}{\Sigma}|)$ defined by

$$L_2(a,|\underset{\sim}{\Sigma}|) = \frac{|\underset{\sim}{\Sigma}|}{a} - \log\frac{|\underset{\sim}{\Sigma}|}{a} - 1 \tag{3.4}$$

it follows that the best equivariant estimator of $|\underset{\sim}{\Sigma}|$ is $c|\underset{\sim}{S}|$ where c minimizes $E_{\underset{\sim}{\Sigma}=I_p}(\frac{1}{c|\underset{\sim}{S}|} + \log c)$. This gives $c = E_{\Sigma=I_p}\{|\underset{\sim}{S}|^{-1}\} = (n-p-2)!/(n-2)!$

As before, we look for a better estimator in the class $\phi(\underset{\sim}{X},\underset{\sim}{S}) = \psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X})|\underset{\sim}{S}|$ for some real valued function ψ. Such a ϕ, under the loss (3.4), has the risk

$$R_\phi = E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}[1/(\psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X})|\underset{\sim}{S}|) + \log\psi(\underset{\sim}{X}^T\underset{\sim}{S}^{-1}\underset{\sim}{X}) + \log|\underset{\sim}{S}| - 1] \tag{3.5}$$

which is minimized for a given $\underset{\sim}{V} = \underset{\sim}{v}$ by choosing $\psi(\underset{\sim}{v}) = \psi_{\underset{\sim}{\xi}_*}(\underset{\sim}{v}) = E_{\underset{\sim}{\xi}_*,I_p}\{|\underset{\sim}{S}|^{-1}|\underset{\sim}{V}=\underset{\sim}{v}\}$. Following Sinha (1976), we can easily show that

$$\psi_{\underset{\sim}{\xi}_*}(\underset{\sim}{v}) \le |I_k+v|\ (n-p-2+k)!/(n-2+k)! = \psi_0(\underset{\sim}{v}) \quad \text{(say)}. \tag{3.6}$$

Then, for every ψ defining $\tilde{\psi}(v) = \min(\psi(v), \psi_0(v))$, it follows that $\tilde{\psi}(X^T S^{-1} X)|S|$ dominates $\psi(X^T S^{-1} X)|S|$ under the loss (3.4). In particular, the testimator $Z = \min\{\frac{(n-p-2)!}{(n-2)!} |S|, \frac{(n-p-2+k)!}{(n-2+k)!}|S + XX^T|\}$ dominates $\{(n-p-2)!/(n-2)!\} |S|$ under L_2 loss.

As in Sinha (1976), it is possible to easily derive sequential testimators of $|\Sigma|$ under both the losses. Details are omitted.

APPENDIX

Proof of Lemma 4. Since, $\underset{\sim}{I}_p + \underset{\sim}{u}\underset{\sim}{u}^T$ is positive definite, there exists a nonsingular $\underset{\sim}{Q}$ such that $\underset{\sim}{Q}\underset{\sim}{Q}^T = \underset{\sim}{I}_p + \underset{\sim}{u}\underset{\sim}{u}^T$. Write $\underset{\sim}{W}_{**} = \underset{\sim}{W}_*\underset{\sim}{Q}$. Then the inequality $\underset{\sim}{\psi}_*(\underset{\sim}{\xi}_*,\underset{\sim}{u}) \le \tilde{\underset{\sim}{\psi}}(\underset{\sim}{u}) \Leftrightarrow \underset{\sim}{\psi}_*^{-1}(\underset{\sim}{\xi}_*,\underset{\sim}{u}) \ge \tilde{\underset{\sim}{\psi}}^{-1}(\underset{\sim}{u})$ can be alternately expressed as $\underset{\sim}{Q}^{T^{-1}} E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}(\underset{\sim}{W}_{**}^T\underset{\sim}{W}_{**}|\underset{\sim}{u})\underset{\sim}{Q}^T \ge (n+k)\underset{\sim}{Q}^{T^{-1}}\underset{\sim}{Q}^T$. Hence, it suffices to show that

$$E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}(\underset{\sim}{W}_{**}^T\underset{\sim}{W}_{**}|\underset{\sim}{u}) \ge (n+k)\underset{\sim}{I}_p \quad \text{for all } \underset{\sim}{u},\underset{\sim}{\xi}_*. \tag{A.1}$$

Note that (A.1) can be alternately expressed as

$$E_{\underset{\sim}{\xi}_*,\underset{\sim}{I}_p}(\underset{\sim}{\ell}^T\underset{\sim}{W}_{**}^T\underset{\sim}{W}_{**}\underset{\sim}{\ell}|\underset{\sim}{u}) \ge (n+k)\underset{\sim}{\ell}^T\underset{\sim}{\ell} \tag{A.2}$$

for all $\underset{\sim}{\ell}(\ne \underset{\sim}{0})$, $\underset{\sim}{u}$ and $\underset{\sim}{\xi}_*$. From (2.19) of Sinha (1976), it follows that a sufficient condition for (A.2) to hold is that

$$\int_{\underset{\sim}{w}\in E^{p^2}} |\underset{\sim}{w}\underset{\sim}{w}^T|^{\frac{n-p+k}{2}} (\underset{\sim}{\ell}^T\underset{\sim}{w}^T\underset{\sim}{w}\underset{\sim}{\ell})\exp[-\tfrac{1}{2}(\underset{\sim}{w}\underset{\sim}{w}^T - 2\underset{\sim}{w}\underset{\sim}{u}_*\underset{\sim}{\xi}_*^T)]d\underset{\sim}{w}$$
$$\div \int_{\underset{\sim}{w}\in E^{p^2}} |\underset{\sim}{w}\underset{\sim}{w}^T|^{\frac{n-p+k}{2}} \exp[-\tfrac{1}{2}(\underset{\sim}{w}\underset{\sim}{w}^T - 2\underset{\sim}{w}\underset{\sim}{u}_*\underset{\sim}{\xi}_*^T)]d\underset{\sim}{w}$$
$$\ge (n+k)\underset{\sim}{\ell}^T\underset{\sim}{\ell} \tag{A.3}$$

for all $\underset{\sim}{\ell} \ne \underset{\sim}{0}$, $\underset{\sim}{u}_* = \underset{\sim}{Q}^{-1}\underset{\sim}{u}$ and $\underset{\sim}{\xi}_*$. In (A.3) and in what follows, we use the notation $\underset{\sim}{w}$ for $\underset{\sim}{w}_{**}$ (and accordingly $\underset{\sim}{W}$ for $\underset{\sim}{W}_{**}$).

Use now the transformation $\underset{\sim}{Z} = \underset{\sim}{W}\underset{\sim}{L}^T$, where $\underset{\sim}{L}^T$ is an orthogonal matrix with its first column vector equal to $\underset{\sim}{\ell}/(\underset{\sim}{\ell}^T\underset{\sim}{\ell})^{1/2}$. We write $\underset{\sim}{Z} = (\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_p)$. Then (A.3) can be alternately expressed as

$$\int_{\underset{\sim}{z}\in E^{p^2}} |\underset{\sim}{z}\underset{\sim}{z}^T|^{\frac{n-p+k}{2}} (\underset{\sim}{z}_1^T\underset{\sim}{z}_1)\exp[-\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{z}\underset{\sim}{z}^T - 2\underset{\sim}{z}\underset{\sim}{u}_L\underset{\sim}{\xi}_*^T)]d\underset{\sim}{z}$$
$$\div \int_{\underset{\sim}{z}\in E^{p^2}} |\underset{\sim}{z}\underset{\sim}{z}^T|^{\frac{n-p+k}{2}} \exp[-\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{z}\underset{\sim}{z}^T - 2\underset{\sim}{z}\underset{\sim}{u}_L\underset{\sim}{\xi}_*^T)]d\underset{\sim}{z}$$
$$\ge n+k, \quad \text{where } \underset{\sim}{u}_L = \underset{\sim}{L}\underset{\sim}{u}_*. \tag{A.4}$$

Next, as in Sinha (1976), let $A_p = I_p$ and

$$A_i = I_p - (Z_{i+1} \cdots Z_p) \begin{pmatrix} Z_{i+1}^T Z_{i+1} & \cdots & Z_{i+1}^T Z_p \\ Z_p^T Z_{i+1} & \cdots & Z_p^T Z_p \end{pmatrix}^{-1} \begin{pmatrix} Z_{i+1} \\ \vdots \\ Z_p \end{pmatrix}$$

$$i = 1, \ldots, p-1.$$

Then, following (2.20)-(2.21) of Sinha (1976) and noting that his w's are our z's, we can express the left hand side of (A.4), in Sinha's (1976) notation, as

$$E[(W_p^T W_p)^{\frac{n-p+k}{2}} (W_{p-1}^T A_{p-1} W_{p-1})^{\frac{n-p+k}{2}} \cdots (W_1^T A_1 W_1)^{\frac{n-p+k}{2}} W_1^T W_1]$$

$$\div E[(W_p^T W_p)^{\frac{n-p+k}{2}} (W_{p-1}^T A_{p-1} W_{p-1})^{\frac{n-p+k}{2}} \cdots (W_1^T A_1 W_1)^{\frac{n-p+k}{2}}] \tag{A.5}$$

where, given $W_{(i+1)}, \ldots, W_{(p)}$, $W^T A_i W_i$ is a noncentral chisquared variable with i.d.f. and noncentrality parameter $\lambda_{(i)}^2 = \eta'_{(i)} A_i \eta_{(i)}$ where $(\eta_{(1)}, \ldots, \eta_{(p)}) = u_L \xi_*^T$. Accordingly, using the fact that $W_1^T W_1 \geq W_1^T A_1 W_1$ and proceeding as in Sinha (1976), we get the expression in (A.5) $\geq 2\,(\frac{n-p+k}{2}) + p = n+k$, where in the ultimate step, one uses (2.22) of Sinha with $r = (n-p+k)/2$. The proof of Lemma 4 is complete.

<u>Proof of Lemma 5</u>. It suffices to show that

$$E[\ell^T (W_*^T W_*)^{-1} \ell \;|u] \leq (n-p-1+k)^{-1} \ell^T (I_p + uu^T) \ell \tag{A.6}$$

for all $\ell (\neq 0)$, u and ξ_*. Defining Q as in the proof of Lemma 4, and using calculations similar to (A.1)-(A.3), we find that (A.6) can be equivalently expressed as

$$\int_{w \in E^{p^2}} |ww^T|^{\frac{n-p+k}{2}} (\ell^T (w^T w)^{-1} \ell) \exp[-\tfrac{1}{2} \mathrm{tr}(ww^T - 2wu_* \xi_*^T)] dw$$

$$\div \int_{w \in E^{p^2}} |ww^T|^{\frac{n-p+k}{2}} \exp[-\tfrac{1}{2} \mathrm{tr}(ww^T - 2wu_* \xi_*^T)] dw$$

$$\leq (n-p-1+k)^{-1} (\ell^T \ell) \tag{A.7}$$

for all $\underset{\sim}{\ell}(\neq \underset{\sim}{0})$, $\underset{\sim}{u}_* = Q^{-1}\underset{\sim}{u}$ and $\underset{\sim}{\xi}_*$. Next make the transformation $\underset{\sim}{Z}^T = \underset{\sim}{W}\underset{\sim}{L}$ where $\underset{\sim}{Z} = (\underset{\sim}{Z}_1,\ldots,\underset{\sim}{Z}_p)$ and $\underset{\sim}{L}$ is an orthogonal matrix with its first column vector equal to $\underset{\sim}{\ell}/(\underset{\sim}{\ell}^T\underset{\sim}{\ell})^{1/2}$. Then, $(\underset{\sim}{Z}\underset{\sim}{Z}^T)^{-1} = (\underset{\sim}{L}^T\underset{\sim}{W}^T\underset{\sim}{W}\underset{\sim}{L})^{-1} = \underset{\sim}{L}^T(\underset{\sim}{W}^T\underset{\sim}{W})^{-1}\underset{\sim}{L}$ so that $\underset{\sim}{\ell}^T(\underset{\sim}{W}^T\underset{\sim}{W})^{-1}\underset{\sim}{\ell}/(\underset{\sim}{\ell}^T\underset{\sim}{\ell})$ is the element in the first row and first column of $(\underset{\sim}{Z}\underset{\sim}{Z}^T)^{-1}$. We denote this by $(\underset{\sim}{Z}\underset{\sim}{Z}^T)^{-1}_{1,1}$. Now, writing $\underset{\sim}{u}_L = \underset{\sim}{L}^T\underset{\sim}{u}_*$, (A.7) can be equivalently expressed as

$$\int_{z\in E^{p^2}} |\underset{\sim}{z}\underset{\sim}{z}^T|^{\frac{n-p+k}{2}} (\underset{\sim}{z}\underset{\sim}{z}^T)^{-1}_{1,1}\exp[-\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{z}\underset{\sim}{z}^T-2\underset{\sim}{z}^T\underset{\sim}{u}_L\underset{\sim}{\xi}_*^T)]dz$$

$$\div \int_{z\in E^{p^2}} |\underset{\sim}{z}\underset{\sim}{z}^T|^{\frac{n-p+k}{2}} \exp[-\tfrac{1}{2}(\underset{\sim}{z}\underset{\sim}{z}^T-2\underset{\sim}{z}^T\underset{\sim}{u}_L\underset{\sim}{\xi}_*^T)]dz$$

$$\leq (n-p-1+k)^{-1}. \tag{A.8}$$

With the same $\underset{\sim}{A}_i$'s $(i=1,\ldots,p)$ as defined in Lemma 4, it follows that

$$|\underset{\sim}{Z}\underset{\sim}{Z}^T| = (\underset{\sim}{Z}_p^T\underset{\sim}{Z}_p)(\underset{\sim}{Z}_{p-1}^T\underset{\sim}{A}_{p-1}\underset{\sim}{Z}_{p-1})\ldots(\underset{\sim}{Z}_2^T\underset{\sim}{A}_2\underset{\sim}{Z}_2)(\underset{\sim}{Z}_1^T\underset{\sim}{A}_1\underset{\sim}{Z}_1)$$

and

$$(\underset{\sim}{Z}\underset{\sim}{Z}^T)^{-1}_{1,1} = |\underset{\sim}{Z}\underset{\sim}{Z}^T|^{-1}(\underset{\sim}{Z}_p^T\underset{\sim}{Z}_p)\ \ldots\ (\underset{\sim}{Z}_2^T\underset{\sim}{A}_2\underset{\sim}{Z}_2).$$

Accordingly, writing $r = (n-p+k)/2$, lhs of (A.8)

$$= E[(\underset{\sim}{Z}_p^T\underset{\sim}{A}_p\underset{\sim}{A}_p)^r\ \ldots\ (\underset{\sim}{Z}_2^T\underset{\sim}{A}_2\underset{\sim}{Z}_2)^r\ (\underset{\sim}{Z}_1^T\underset{\sim}{A}_1\underset{\sim}{Z}_1)^{r-1}]$$

$$\div E[(\underset{\sim}{Z}_p^T\underset{\sim}{A}_p\underset{\sim}{Z}_p)^r\ \ldots\ (\underset{\sim}{Z}_2^T\underset{\sim}{A}_2\underset{\sim}{Z}_2)^r(\underset{\sim}{Z}_1^T\underset{\sim}{A}_1\underset{\sim}{Z}_1)^r]$$

$\leq (2r-1)^{-1}$(using conditional argument as in the proof of Lemma 4 and (2.22) of Sinha (1976))

$= (n-p+k-1)^{-1}$.

The proof of Lemma 5 is complete.

REFERENCES

[1] BREWSTER, J.F. and ZIDEK, J.V. (1974). Improving on equivariant estimators. Ann. Statist. 2, 21-38.

[2] DEY, D.K., and SRINIVASAN, C. (1985). Estimation of covariance matrix under Stein's loss. To appear in Ann. Statist.

[3] GIRSHICK, M.A. and SAVAGE, L.J. (1951). Bayes and minimax estimates for quadratic functions. Proc. of the 2nd Berk. Symp. on Math. Stat. and Prob. Univ. California, Berkeley, 53-73.

[4] HAFF,L.R. (1977). Minimax estimators for a multinormal precision matrix. J. Mult. Anal., 7, 374-385.

[5] HAFF, L.R. (1979a). An identity for the Wishart distribution with applications. J. Mult. Anal., 9, 531-542.

[6] HAFF, L.R. (1979b). Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity. Ann. Statist., 7, 1264-1276.

[7] HAFF, L.R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. Ann. Statist. 8, 586-597.

[8] HAFF, L.R. (1982). Solutions of the Euler-Lagrange equations for certain multivariate normal estimation problems. Submitted.

[9] HODGES, J.L., Jr., and LEHMANN, E.L. (1951). Some applications of the Cramer-Rao inequality. Proc. 2nd Berk. Symp. Math. Statist. and Prob. Univ. of California, Berkeley, 13-22.

[10] JAMES, W., and STEIN, C. (1961). Estimation with quadratic loss. Proc. 4th Berk. Symp. Math. Statist. and Prob. Univ. of California, Berkeley V1, 361-379.

[11] SHORROCK, R.W. and ZIDEK, J.V. (1976). An improved estimator of the generalized variance. Ann. Statist., 4, 629-638.

[12] SINHA, B.K. (1976). On improved estimators of the generalized variance. J. Mult. Anal., 6, 617-625.

[13] STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Stat. Math., 16, 155-160.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 85-0698

2. GOVT ACCESSION NO. A159 059

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Inadmissibility of the best equivariant estimators of the variance-covariance matrix and the generalized variance under entropy loss.

5. TYPE OF REPORT & PERIOD COVERED
Technical - July 1985

6. PERFORMING ORG. REPORT NUMBER
TR- 85-27

7. AUTHOR(s)
Bimal K. Sinha and Malay Ghosh

8. CONTRACT OR GRANT NUMBER(s)
F49620-85-C-0008 [redacted]

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Center for Multivariate Analysis
515 Thackeray Hall
Univeristy of Pittsburgh, Pittsburgh, PA 15260

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
61102F 2304 A5

11. CONTROLLING OFFICE NAME AND ADDRESS
Air Force Office of Scientific Research
Department of the Air Force
Bolling Air Force DC 20332

12. REPORT DATE.
July 1985

13. NUMBER OF PAGES
18

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Best equivariant estimator, entropy loss, generalized variance, MANOVA test, Roy's maximum root test, testimator, Wishart distribution.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Based on a data matrix $X = (\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_k)$: $p \times k$ with independent columns $\underset{\sim}{X}_i \sim N_p(\underset{\sim}{\xi}_i, \Sigma)$, and an independent Wishart matrix S: $p \times p \sim W_p(n, \Sigma)$, testimators dominating the best equivariant estimators of Σ and $|\Sigma|$ are obtained under two types of entropy loss. For simultaneous estimation of the mean vector and the variance covariance matrix of a multinormal population, a suitable entropy loss is developed and testimators dominating the pair consisting of the sample mean vector and the best multiple of the sample Wishart matrix are derived. A technique of

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Sinha (Jour. Mult. Analysis, 1976) is heavily exploited.

END

FILMED

11-85

DTIC